

intersession interval can be evaluated adequately.

REFERENCES

COTTON, J. W., & LEWIS, D. J. Effects of intertrial interval on acquisition and

extinction of a running response. *Journal of Experimental Psychology*, 1957, 54, 15-20.

DENNY, M. R. Relaxation theory and experiments. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press, in press.

HERRNSTEIN, R. J. Method and theory in

the study of avoidance. *Psychological Review*, 1969, 76, 49-69.

LEVIS, D. J. Between-trial transporting of animals: A methodological consideration. *Behavior Research Methods & Instrumentation*, 1970, 2, 157-160.

LOGAN, F. A. *Incentive*. New Haven: Yale University Press, 1960.

## Measuring clustering differences among categories in the same list\*

KARL HUNT

University of Tennessee at Chattanooga, Chattanooga, Tennessee 37401

A method is presented for analysis of clustering differences among categories in the same list. This allows category characteristics to be varied within Ss by using mixed lists, whereas previously only homogeneous lists could be employed.

When a list of items is presented for free recall, related items are often recalled contiguously, regardless of their order of presentation. Bousfield (1953) called this phenomenon "clustering" and measured it by counting repetitions. A repetition is the contiguous recall of two related items; thus a sequence of three such items contains two repetitions. For example, S might recall, "... horse, oak, pine, maple, book..." If the relationship of interest is "tree-ness," the pairs oak-pine and pine-maple each represent a repetition. If "birch" were recalled in place of "book," the number of repetitions would increase from two to three. Various types of relationships have been used in studies of clustering, e.g., membership in the same taxonomic category (Bousfield & Puff, 1965), common grammatical classification (Koplin & Moates, 1968), phonemic similarity (Bousfield & Wicklund, 1969), etc. Here the term *category* will be arbitrarily used to refer to any group of related items whatever the relationship may be.

In a randomly ordered list of items, a certain number of repetitions may occur by chance. Therefore, some repetitions in a response protocol might reflect chance ordering rather than S's recall processes. Bousfield and Bousfield (1966) provide equations for the expected values of chance repetitions for an entire list, and for individual categories within the list:

$$E(R) = \left( \sum_{i=1}^k m_i^2 / n \right) - 1 \quad (1)$$

$$E(R_i) = m_i(m_i - 1) / n \quad (2)$$

where R is the number of repetitions

for the whole list,  $R_i$  the repetitions for Category i,  $m_i$  the number of items recalled from Category i, n the total number of items recalled, and k the number of categories represented by items in the list. By subtracting the expected chance values from observed values, one obtains a measure of the extent to which S has actually clustered the items. The measure is designed to be independent of the number of items recalled, as n and m are treated as constants for a given recall sequence.

Unfortunately, these equations are appropriate only for measuring clustering for an entire list rather than separate categories within the list. As Bousfield and Bousfield indicate, Eq. 2 does not take into account the amount of clustering in categories other than i. Because more clustering in the other categories effectively increases  $E(R_i)$  beyond the value given by Eq. 2, it cannot be used to compare separate categories or types of categories within lists. An additional problem, highly similar to that discussed by Hudson and Dunn (1969), is that the variance for the distribution of chance values of  $R_i$  will change with different values of n, R, and  $m_i$ . Again, this prevents comparison of intralist categories with the Bousfield and Bousfield formula. The methods presented here correct these problems, allowing estimation of clustering differences among the categories in a single stimulus list.

### REMOVAL OF INTERCATEGORY CLUSTERING DEPENDENCIES

Equations 1 and 2 assume that the n elements of a list are independently ordered, i.e., in a random sequence. However, if S recalls two items together because of some common feature, the two items are no longer independently ordered. Therefore, they should be treated as a single item

rather than two. Why this is so might be clarified by a rearrangement and examination of Eq. 2:

$$E(R_i) = (m_i - 1)m_i/n \quad (2a)$$

As stated before, n and  $m_i$  are fixed for a given recall protocol. The S's behavior is analogous to placing all n items recalled into a row of n boxes, one item per box, where  $m_i$  of the boxes will hold the Category i items. The first factor in Eq. 2a,  $m_i - 1$ , is the number of boxes in which a Category i item may be placed adjacent to another box already holding a Category i item. The 1 is subtracted since additional category members cannot be placed next to the last item in the category. The second factor,  $m_i/n$ , is the probability that any one of the n boxes will be filled by one of the  $m_i$  items, assuming random selection. Multiplying the two factors is the same as adding the probabilities for the  $(m_i - 1)$  positions. If the number of boxes becomes less than n, then there are fewer boxes that can be filled by an  $m_i$  item. In turn,  $m_i/n$  will underestimate the probability of an item being in a particular box.

Clustering of two items from a non-i category has the effect of removing one of the n boxes available to Category i items, because the second item in the cluster must, by definition, occupy the box adjacent to the first. A possible solution to the problem might be to subtract the number of repetitions in non-i categories from n. However, this would be an overcorrection if some of the repetitions occurred by chance. Instead, only those repetitions which exceed chance should be subtracted. If Eqs. 1 and 2 are used to calculate the chance repetitions, the corrected value of  $E(R_i)$  is given by

$$E(R_i)' = \frac{m_i(m_i - 1)}{n_i'} \quad (3)$$

where

$$n_i' = n - [R - E(R)] + [R_i - E(R_i)] \quad (4)$$

The first term in brackets in Eq. 4 is, of course, the number of "nonchance" repetitions for an entire list and is a

\*Supported in part by a Faculty Research Grant from the University of Tennessee at Chattanooga.

Table 1  
Sample Response Protocols

S	Items
1	bluejay, cardinal, creme, lotion, make-up, ruby, kaolin, galena, China, Russia, England, France, actor, football, golf, tennis, Vermont, Kansas, Maine, rolls, rye, wheat
2	sparrow, cardinal, robin, France, England, China, football, bluejay, tennis, golf, lotion, make-up, creme, Kansas, Maine, Utah, actor, rye, wheat, Russia, chemist, soap

constant for all categories in the list. The addition to the left side of the second term in brackets has the effect of removing Category *i* from the correction factor. Also, note that Eq. 3 reduces to Eq. 2 when the null hypothesis holds for all of the non-*i* categories, i.e., when

$$R - R_i = E(R) - E(R_i)$$

STANDARDIZATION

Hudson and Dunn (1969) have observed that the variance of chance values of *R* depends on the number of categories (*k*), and number of items (*m*<sub>1</sub>, ..., *m*<sub>*k*</sub>) from each category. They suggest dividing the difference between *R* and *E*(*R*) by the standard deviation of *E*(*R*) in order to equate clustering scores for different *S*s. Since the chance distribution of *R*<sub>*i*</sub> varies with *n* and *m*<sub>*i*</sub>, the clustering index for Category *i* should be standardized as well:

$$SCI_i = \frac{R_i - E(R_i)'}{SD[E(R_i)]} \quad (5)$$

where *SCI*<sub>*i*</sub> is the standardized clustering index for Category *i*, and the standard deviation for the chance distribution of *R*<sub>*i*</sub> is given by

$$SD[E(R_i)]$$

$$= \left[ \frac{m_i(m_i - 1)(n_i' - m_i)(n_i' - m_i + 1)}{(n_i')^2(n_i' - 1)} \right]^{1/2}$$

(6)

Equation 6 is based on the multiple-runs distribution as discussed by David and Barton (1962).

COMMENTS

Table 1 provides sample response protocols for two *S*s, while Table 2 illustrates the application of Eqs. 2-6 to the categories in the protocols. The protocols have the same *n* (22) but different *R*s (14 for the first *S* and 10 for the second), thus showing how the total amount of clustering in a list affects the clustering scores for individual categories despite equal numbers of items recalled. The first column in Table 2 gives estimates of clustering when Eq. 2 is used to calculate the amount of chance clustering. The next column shows how *n*<sub>*i*</sub>', the corrected value of *n* for Category *i* as given by Eq. 4, will vary with the amount of clustering in non-*i* categories. The third column gives new estimates of clustering calculated by Eq. 3. These values are generally smaller than the corresponding items in Column 1, since adjustment has been made for the underestimation of chance clustering made by Eq. 2. Column 4 gives the appropriate standard deviations calculated by Eq. 6. In turn, these are divided into the Column 3 values, as indicated by Eq. 5, yielding the *SCI*<sub>*i*</sub> scores in the last column.

While the present procedures will affect the relationships among clustering estimates for categories within the same *S*'s protocol, the greatest changes occur between *S*s. For example, the first estimate of clustering for "states" was 1.73 for both *S*s, as *m*<sub>*i*</sub> and *R*<sub>*i*</sub> were the same for both. The final estimates or *SCI*s, however, differed by 0.53 standard units.

It can be seen in Table 2 that no calculations were performed for two categories, "professions" for the first *S* and "minerals" for the second. Table 1 shows that the first *S* recalled only one profession (actor) and that the second *S* did not recall any minerals, although both *S*s were given the same stimulus list. Since the procedures described here are designed to measure clustering separately from recall (i.e., when enough items are recalled for clustering to take place), they cannot be applied to categories with *m*<sub>*i*</sub>s less than 2. The exclusion of these categories from an analysis of clustering could result in an unequal number of categories representing various intralist conditions. For example, *E* may present, in the same stimulus list, familiar items from some categories and unfamiliar items from other categories. If fewer unfamiliar items are recalled, their categories are less likely to be represented in the clustering analysis. An adjustment must be made for the unequal frequencies if a statistical analysis is not to reflect these differences in recall. The procedure used by this investigator is to replace the missing categories with artificial data according to a method devised by Yates and presented in Cochran and Cox (1957).

It should be emphasized that the *metric* independence of clustering and recall does not imply *psychological* independence. The present procedure is designed to measure only relationships among items actually recalled, not between those recalled and those not recalled. The fact that *S* recalls a single item from a category rather than none at all could carry implications for the psychological relationship between clustering and recall. However, there is no way of telling if the failure to recall a second item corresponds to a failure of the processes responsible for clustering or to other recall processes not related to clustering. This sort of question would have to be answered with methods other than those presented here.

Table 2  
Calculated Values for Sample Protocols

S	Category	<i>R</i> <sub><i>i</i></sub> - <i>E</i> ( <i>R</i> <sub><i>i</i></sub> )	<i>n</i> <sub><i>i</i></sub> '	<i>R</i> <sub><i>i</i></sub> - <i>E</i> ( <i>R</i> <sub><i>i</i></sub> )'	SD[ <i>E</i> ( <i>R</i> <sub><i>i</i></sub> )]	<i>SCI</i> <sub><i>i</i></sub>
1	Birds	0.91	10.91	0.82	0.39	2.11
	Cosmetics	1.73	11.73	1.49	0.59	2.53
	Minerals	1.73	11.73	1.49	0.59	2.53
	Nations	2.45	12.46	2.07	0.73	2.77
	Professions					
	Sports	1.73	11.73	1.49	0.59	2.53
	States	1.73	11.73	1.49	0.59	2.53
	Types of Bread	1.73	11.73	1.49	0.59	2.53
2	Birds	1.45	15.82	1.24	0.70	1.77
	Cosmetics	1.45	15.82	1.24	0.70	1.77
	Minerals					
	Nations	1.45	15.82	1.24	0.70	1.77
	Professions	-0.09	14.27	-0.14	0.35	-0.40
	Sports	0.73	15.09	0.60	0.54	1.11
	States	1.73	16.09	1.63	0.53	3.06
	Types of Bread	0.91	15.27	0.87	0.34	2.58

REFERENCES

BOUSFIELD, W. A. The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, 1953, 49, 229-240.  
 BOUSFIELD, A. K., & BOUSFIELD, W. A. Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports*, 1966, 19, 935-942.  
 BOUSFIELD, W. A., & PUFF, C. R. Determinants of the clustering of taxonomically and associatively related word pairs. *Journal of General Psychology*, 1965, 73, 211-221.  
 BOUSFIELD, W. A., & WICKLUND, D. A. Rhyme as a determinant of clustering. *Psychonomic Science*, 1969, 16, 183-184.  
 COCHRAN, W. G., & COX, G. M.

- Experimental designs.* (2nd ed.) New York: Wiley, 1957.
- DAVID, F. N., & BARTON, D. E. *Combinatorial chance.* New York: Hafner, 1962.
- HUDSON, R. L., & DUNN, J. E. A major modification of the Bousfield and Bousfield (1966) measure of category clustering. *Behavior Research Methods & Instrumentation*, 1969, 1, 110-111.
- KOPLIN, J. H., & MOATES, D. R. Form-class clustering in recall with relevant or irrelevant prior exposure. *Journal of Verbal Learning & Verbal Behavior*, 1968, 7, 263-264.